ehp ENVIRONMENTAL HEALTH PERSPECTIVES
http://www.ehponline.org

Scientific Considerations for Evaluating Cancer Bioassays Conducted by the Ramazzini Institute

Jeffrey S. Gift, Jane C. Caldwell, Jennifer Jinot, Marina V. Evans, Ila Cote, and John J. Vandenberg

http://dx.doi.org/10.1289/ehp.1306661

Received: 16 February 2013
Accepted: 13 September 2013
Advance Publication: 17 September 2013

NIH National Institute of Environmental Health Sciences

Scientific Considerations for Evaluating Cancer Bioassays Conducted by the Ramazzini Institute

Jeffrey S. Gift,[1] Jane C. Caldwell,[1] Jennifer Jinot,[1] Marina V. Evans,[2] Ila Cote,[1] and John J. Vandenberg[1]

[1]United States Environmental Protection Agency, National Center for Environmental Assessment, Research Triangle Park, North Carolina, USA

[2]United States Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Research Triangle Park, North Carolina, USA

Corresponding Author:

Jeff Gift, Ph.D.

U.S. EPA, National Center for Environmental Assessment

109 TW Alexander Dr

Mail Code B243-01

Research Triangle Park, NC 27711

United States

919-541-4828

gift.jeff@epa.gov

Short title: Considerations for Ramazzini Institute Bioassays

Acknowledgements:

We thank Linda Birnbaum (NIEHS/NTP; formerly at EPA), Lyle Burgoon (EPA), John Cowden (EPA), J. Allen Davis (EPA), Danielle Devoney (EPA), Reeder Sams (EPA), and David

Svendsgaard (EPA) for their manuscript reviews and suggestions, and the EPA Health and Environmental Research Online (HERO) (U.S. EPA 2009) team. References cited in this document, including bibliographic information and abstracts, can be found at http://hero.epa.gov/ramazzini.

Financial Disclosure Declaration: The authors declare that they have no actual or potential competing financial interests.

Disclaimer: This manuscript has been reviewed by the U.S. EPA and approved for publication. The views expressed in this manuscript are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA.

ABSTRACT

Background: The Ramazzini Institute (RI) has completed nearly 400 cancer bioassays on over 200 compounds. Study design and protocol differences between RI and other laboratories have been suggested by the European Food Safety Authority (EFSA) and others to contribute to controversy regarding cancer hazard findings, principally lymphoma/leukemia diagnoses.

Objective: Evaluate RI study design, protocol differences, and accuracy of tumor diagnoses for their impact on carcinogenic hazard characterization.

Methods: We analyzed the findings from a recent Pathology Working Group (PWG) review of RI procedures and tumor diagnoses, evaluated consistency of RI and other laboratory findings for chemicals identified by RI as positive for lymphoma/leukemia, and examined evidence for a number of other issues raised regarding RI bioassays. The RI cancer bioassay design and protocols were evaluated in the context of relevant risk assessment guidance from International authorities.

Discussion: Although the PWG identified close agreement with RI diagnoses for most tumor types, it did not for lymphoma/leukemia of the respiratory tract and neoplasms of the inner ear and cranium. We discuss (1) the implications of the PWG findings, particularly lymphoma diagnostic issues, (2) differences between RI and other laboratory studies that are relevant to evaluating RI cancer bioassays and (3) future work that may help resolve some concerns.

Conclusions: We conclude that (1) issues related to respiratory tract infections have complicated diagnoses at that site (i.e., lymphoma/leukemia) and for neoplasms of the inner ear and cranium, and (2) there is consistency and value in RI studies for identification of other chemical-related neoplasia.

INTRODUCTION

The Ramazzini Institute (RI) is an independent non-profit organization that has been conducting lifetime cancer bioassays in rodents since 1970. RI's stated approach for their studies includes: (1) "Use of animal species and strains whose basic tumorigram and kind of response to cancer stimuli is not too remote from the human counterpart;" (2) "Continuing bioassays until the end of the life of an animal;" (3) "Following the rules of Good Laboratory Practice as a minimum standard in experiment management;" (4) "Choosing precise parameters to assess neoplastic response;" and (5) "Standardizing the experimental conditions for conducting experiments, parameter assessment, and data presentation" (Maltoni et al. 1999).

Perceived problems in RI studies have been central in European Food Safety Authority (EFSA) reviews (EFSA 2006, 2009) of RI aspartame bioassay findings of lymphoreticular tumors (Soffritti et al. 2005; Soffritti et al. 2006b; Soffritti et al. 2007). The level of inflammatory changes in the lungs of RI animals has prompted discussions regarding the role of respiratory infections in tumor formation, the ability to discern tumors from inflammatory infiltrates, and the adequacy of RI protocols (Caldwell et al. 2008; Cruzan 2009; EFSA 2006, 2009; NTP 2011; Schoeb et al. 2009; Schoeb and Mcconnell 2011a, b).

This review summarizes: (1) recent U.S. Environmental Protection Agency (EPA) and National Toxicology Program (NTP) efforts to investigate the issues raised by EFSA and others related to RI chronic bioassays; (2) relevant considerations for RI cancer bioassays evaluation that take into account the unique aspects of the RI study design and protocols in the context of existing international risk assessment guidelines; and (3) methods and approaches that may assist in the future conduct and review of RI chronic bioassays.

INVESTIGATIONS

Pathology Working Group (PWG) Review

In April 2010, pathologists and technicians representing the NTP visited the RI in Bentivoglio, Italy and conducted a preliminary review of RI pathology procedures and lymphoma/leukemia diagnoses from the RI methanol study (Malarkey et al. 2010). In 2011, NTP and EPA sponsored a more comprehensive PWG review by an independent team from Experimental Pathology Laboratories (EPL). The PWG review included select tissues from RI studies of methanol, methyl-tertiary-butyl ether (MTBE), ethyl-tertiary-butyl ether (ETBE), vinyl chloride and acrylonitrile. A summary of the PWG results (NTP 2011) and full Pathology Quality Assessment Review and PWG Coordinators Reports for the five RI studies (EPL 2011a, b, c, d, 2012) are publically available.

As part of the 2011 PWG review, nearly all RI study slides were examined by a quality assessment (QA) pathologist(s) who provided a more complete diagnosis and comparison of all lesions that were initially diagnosed by RI pathologists. A subset of slides of interest for each chemical was then selected for a PWG panel review. The most thorough reviews were for methanol [i.e., 13,011 slides from 800 rats reviewed by 3 QA pathologists (EPL 2011b)] and MTBE [i.e., 6,751 slides from 360 rats reviewed by a single QA pathologist (EPL 2011c)]. The focus of the PWG panel reviews was narrowed to an examination of lymphoma/leukemia and ear/cranium neoplasm diagnoses for methanol (i.e., 744 slides from 367 rats) and lymphoma/leukemia diagnoses and testicular tumor for MTBE (i.e., 179 slides from 74 rats). More limited reviews were conducted of ETBE [i.e., oral cavity, uterus, and vagina (EPL 2011a)], vinyl chloride [i.e., liver tumors (EPL 2011d)], and acrylonitrile [i.e., brain/central

nervous system, extrahepatic angiomatous lesions, zymbal gland, liver and mammary gland (EPL 2012)] slides. As indicated in Table 1, there was general agreement between RI and QA or PWG pathologists for a large number of tumor diagnoses from these five reviews and three other, more limited NTP pathology reviews (Cesta 2008; Hailey 2004; Hailey 2001; Malarkey et al. 2010) of RI study results.

However, a consistent feature of the 2010 preliminary review of the RI methanol study and the 2011 PWG review of RI methanol and MTBE studies has been the difficulty distinguishing between lymphoma/leukemia and ear/cranium neoplasms from concurrent lung infection or inflammatory infiltrates. As noted in the PWG summary report (NTP 2011) and as discussed in Caldwell et al. (2008), there were end-of-life infections present in the lungs of these RI study rats. The 2010 preliminary review noted diagnostic agreement of lymphoma/leukemia when sites outside the lung were affected (Malarkey et al. 2010). As shown in Table 2, although reporting a dose-dependent increase in lymphomas/leukemias in MTBE-treated female rats, the PWG panel did not find treatment-related increases for these tumors in rats treated with methanol. Also, fewer lymphomas/leukemias were diagnosed for both chemicals by the PWG panel than by RI and QA pathologists for all treatment groups. The 2011 PWG report gives a consensus opinion representing a majority of the participants, but also noted that occasional differences of opinion were discussed until a consensus diagnosis was achieved. The diagnostic differences between pathologists in the PWG review of the methanol (EPL 2011b) and MTBE (EPL 2011c) RI studies appear to largely reflect difficulties discerning lymphoma in the lungs of infected rats, but the following factors may have contributed as well.

The accuracy of diagnoses of pathological lesions can be affected by autolysis of tissue cells and the fixation process used to prepare pathology slides (see "*Complete and peer-reviewed*

histopathological evaluations"). However, the full pathology reports for the RI methanol and MTBE studies state that the histological quality of the sections was considered to be "very good by the QA pathologist" and that "neither the occasional cases with tissue autolysis nor the use of alcohol fixation presented diagnostic difficulties" (EPL 2011b, c).

The consistency of diagnoses of lymphoma/leukemias could also have been affected by differing categorization schemes used by RI, QA and PWG pathologists. Lymphomas encompass a spectrum of histological types, and many schemes have been developed to describe them (Harris et al. 1994; Harris et al. 1999; Swerdlow et al. 2008). Updates to the Revised European-American Lymphoma (REAL) scheme report a consensus that, although exhibiting different clinical manifestations, precursor neoplasms (e.g. lymphoblastic lymphomas) presenting as solid tumors or with marrow and blood involvement are biologically the same disease (Harris et al. 1999). In many RI studies, specific histological types (e.g., lymphoblastic lymphoma, lymphoblastic leukemia, lymphocytic lymphoma, lymphoimmunoblastic lymphoma, histiocytic sarcoma, monocytic leukemia, and myeloid leukemia) were collectively referred to as either "hemolymphoreticular neoplasms" or "lymphomas/leukemias" and combined for reporting of their study results. However for RI MTBE studies, incidences of lymphoma subtypes (e.g., lymphoblastic and immunoblastic lymphomas) were reported based on histological examination via light microscopy (Belpoggi et al. 1999). These tumor cell subtype descriptions, especially for immunoblastic tumor cells, were consistent with those reported by others (Frith 1988; Frith et al. 1993; Otová et al. 2002).

RI diagnoses of neoplasms as lymphocytic, histiocytic, monocytic and/or myeloid in origin were generally confirmed in a preliminary NTP pathology review of RI aspartame slides, but the RI's practice of combining myeloid leukemias and histiocytic sarcomas with malignant lymphomas

was not accepted, as these neoplasms are considered to be of different cellular origins (Hailey 2004). EPA (U.S. EPA 2005a) and other investigators (EFSA 2006; McConnell et al. 1986) have also expressed the opinion that tumors of different cellular origins should be treated as separate malignancies and not combined for statistical evaluation. The QA and PWG pathologists defined lymphoreticular neoplasms as either malignant lymphomas or mononuclear cell leukemias so that their results would be more comparable to reporting schemes historically used by NTP and because these tumors are thought to be of separate cellular origin after differentiation from myeloid stem cells (EPL 2011b, c). Therefore, the full methanol and MTBE PWG reports (EPL 2011b, c) do not contain information that is directly comparable to RI study reports in this regard.

Differences in protocols used by RI and the reviewing groups may also have affected diagnostic consistency. Both the 2010 preliminary and 2011 comprehensive reviews typically based lymphoma/leukemia conclusions on the occurrence of the lesions outside the lung (e.g., thymus, spleen, liver, and lymph nodes). The limited number of slides reviewed by the 2011 PWG panel affected the ability to fulfill the requirement of additional sites for a definitive diagnosis. In some cases, the PWG panel reviewed lung lymphomas without also reviewing potentially corroborating diagnoses in other tissues made by QA pathologists. Protocol differences between the methanol and MTBE QA reviews [e.g., three pathologists were used for the methanol QA (EPL 2011b), whereas one pathologist was used for the MTBE QA (EPL 2011c)] provide another possible source of diagnostic variability.

RI Findings Relative to Other Laboratories

Huff (2002) evaluated bioassay results for fourteen chemicals studied by both RI and NTP laboratories and reported consistent carcinogenicity conclusions for eleven chemicals, nine with carcinogenic activity and two without. For xylene, one of the three chemicals with apparent inconsistent findings, NTP and RI tested different mixtures (e.g., NTP's mixture contained 17% ethylbenzene) and study results were not completely discordant (i.e., NTP reported "no evidence" and RI reported "non-dose related" evidence of carcinogenic activity). Vinylidene chloride and toluene, the two other chemicals with discordant results, were tested via different routes of exposure (i.e., toluene exposure was inhalation by NTP and gavage by RI; vinylidene chloride exposure was gavage by NTP and inhalation by RI). Also, Huff (2002) reported that the positive RI findings for toluene were "less than overwhelming," the negative NTP findings for vinylidene chloride "…less than adequate because the use of a maximum tolerated dose had not been clearly demonstrated," and the positive RI findings for vinylidene chloride were for "…increases in leukemias and total malignant tumors in Sprague-Dawley rats whose exposure began *in utero*." Thus, the Huff (2002) review indicates a general consistency between RI and NTP for the identification of carcinogenic agents, with differences in chemical purity and study design being possible explanations for discordant results. Given the difficulties and recent controversy associated with the diagnosis of lymphomas/leukemias in RI studies, we performed an analysis to determine whether the positive RI findings for this endpoint are consistent with the results of other laboratories.

Of over 200 compounds tested (Soffritti et al. 2002a), RI has reported dose-related increases in the incidence of lymphomas/leukemias for ten (i.e., aspartame, chlorinated drinking water, di-isopropyl-ether [DIPE], formaldehyde, mancozeb, methanol, MTBE, tert-amyl-methyl-ether

[TAME], toluene, and vinylidene chloride). The findings of RI and non-RI cancer bioassays for these ten lymphoma/leukemia–positive RI chemicals are summarized in Table 3. Only RI performed cancer bioassays for DIPE, mancozeb and TAME. For the seven chemicals studied by both RI (nine total studies) and non-RI (nineteen total studies) laboratories, three (i.e., chlorinated drinking water, methanol, and MTBE) have been reported to be positive for lymphomas/leukemias in non-RI laboratories. These findings include: (1) marginal increases in leukemias in female rats exposed to chlorinated drinking water (NTP 1992); (2) positive findings in Eppley Swiss Webster mice exposed to methanol (Apaja 1980); and (3) an increase in mononuclear cell leukemia in rats co-exposed to MTBE and gasoline, but not gasoline alone [reported by Burns and Melnick (2012) using data from Benson et al. (2011)].

Dissimilar study results may be attributable not only to pathology diagnostic issues discussed previously, but study design differences as well. Important design differences across laboratories include overall study duration, exposure route, and species/strain. Only one non-RI bioassay used an RI-like lifespan protocol (i.e., 160 weeks) (Apaja 1980), and only two employed the same species and strain used by RI (Molinary 1984; NTP 1982). Similar routes of exposure were only used for studies of aspartame, chlorinated drinking water, and formaldehyde. Furthermore, only RI studies started exposures *in utero* (i.e., for vinylidine chloride, Cotti et al. 1988; and for aspartame, Soffritti et al. 2007).

In support of a chemical relationship for increased lymphoma/leukemia incidence, Soffritti et al. (2006b) noted that several positive RI studies for this endpoint involved either formaldehyde or chemicals that are metabolized to formaldehyde (i.e., methanol, aspartame, and MTBE). Consistency of liver tumor induction in rodents exposed to these same metabolized compounds support the plausibility of such a linkage (BushyRun 1992a; Soffritti et al. 2002a; Soffritti et al.

2010). Formaldehyde has been classified as a "human carcinogen" (Cogliano et al. 2005), and has been causally associated with leukemia (Baan et al. 2009; Zhang et al. 2009) and possibly lymphomas (Zhang et al. 2009) in humans.

Further, a simplistic examination of tumor site concordance using differing study designs may not capture mechanistic concordance and susceptibility differences that can indicate carcinogenic potential. Other tumor types have been observed in non-RI studies for methanol [lung tumors and pheochromocytomas in rats, (NEDO 1985)], formaldehyde [nasal cavity squamous cell carcinomas in rats (Kerns et al. 1983)], MTBE [hepatocellular tumors in mice (BushyRun 1992a); Leydig interstitial cell adenomas, brain astrocytomas, and renal tubule tumors in rats (Bermudez et al. 2012; Burns and Melnick 2012; BushyRun 1992a, b)], and MTBE in gasoline vs. gasoline alone [renal tubule tumors and Leydig interstitial cell adenomas in rats (Benson et al. 2011; Burns and Melnick 2012)].

The diagnosis of increased lymphomas/leukemias in a minority of RI studies (i.e., ~5%) and consistency of diagnoses between RI and non-RI studies for some chemicals (especially those metabolized to formaldehyde) suggests that a regular misassociation of the endpoint and chemical exposures has not occurred in RI studies. However, the lack of evidence of treatment-related increases of lymphomas/leukemias reported in non-RI studies for 7 of the 10 chemicals listed in Table 3 indicates that risk assessors should carefully consider all possible explanations for the lack of tumor site concordance, including differences in study design and laboratory protocols.

CONSIDERATIONS FOR EXISTING RI CANCER BIOASSAYS

Guidance and Study Design Criteria

In order to identify the most relevant considerations for evaluations of RI cancer bioassay design and protocols, we reviewed existing International guidelines from a number of sources. Bioassay design issues are discussed in guidelines by the NTP (Melnick et al. 2008; NTP 2006), in EPA pesticide program guidance (U.S. EPA 1998), and in FDA "Redbook" guidelines (Hinton 2000). Acceptable quality assurance procedures or Good Laboratory Practices (GLP) guidelines are also available (Lilly et al. 1995; OECD 2007). The interpretation of bioassay results is the primary focus of EPA's carcinogen risk assessment guidelines (2005a, b). Guidance on the process of peer review and evaluation of cancer bioassays are provided by EPA (U.S. EPA 2006) and the International Agency for Research on Cancer (IARC 2006).

Keeping in mind current standards as well as those contemporaneous with the study, EPA carcinogenic risk guidelines (U.S. EPA 2005a) encourage the use of established criteria [e.g., NTP guidelines (NTP 2006)] for judging the technical adequacy of individual animal carcinogenicity studies. We have used NIEHS study design considerations identified in the comprehensive review by Melnick et al. (2008) to analyze RI study design, protocols and reporting (i.e., use of: sensitive animal models for endpoints under investigation; detailed characterization of the agent and administered doses; challenging doses and durations of exposure and observation; sufficient numbers of animals per dose group; multiple dose groups for characterization of dose-response relationships; complete and peer-reviewed histopathological evaluations; and pairwise comparisons and analyses of trends based on

survival-adjusted incidence). These considerations are also cross-referenced in the guidelines noted above.

Animal models that are sensitive for endpoints under investigation

In the early 1970s, the RI and the National Cancer Institute (NCI) used Sprague-Dawley rats in their cancer bioassays; by the late 1970s, NCI (and other laboratories, including NTP) switched to Fischer 344 N (F344N) rats. The RI did not switch strains and the U.S. Food and Drug Administration (FDA) still primarily uses the Sprague-Dawley strain to assess the effects and safety of drugs and additives (Duffy et al. 2008). In 2009, NTP started to transition back to Harlan Sprague-Dawley rats for its cancer bioassays (King-Herbert et al. 2010) because of health-related concerns for the F344N colony (e.g., a high incidence of leukemia and Leydig cell tumors, declining fertility, sporadic seizures, and chylothorax).

The historical databases for RI and NTP studies reflect differences in rat strain sensitivity and ability to detect certain types of cancer (e.g., prostate tumors and leukemias) (Melnick et al. 2008). Such differences have implications for comparisons and interpretation of bioassay data. FDA (Hinton 2000) recommends that new drug applicants consider "…the responsiveness of particular organs and tissues" in addition to general sensitivity when selecting rodent species, strains, and substrains for testing.

Laboratory animal and human cancers do not always occur in analogous or the same target/system [e.g., rodent Zymbal gland tumors were the first and most consistent benzene-induced cancer response observed, although humans do not possess these glands]. Such rodent cancer findings should not be dismissed given that growth control mechanisms at the cellular level are generally homologous among mammals (U.S. EPA 2005a). Coherence of tumor

induction, but not necessarily tumor site concordance, across species may reflect similarities in metabolism, cell signaling perturbations, and cancer susceptibility despite differing species/strain/gender sensitivity or study design.

With regard to lymphoma/leukemia diagnoses, the types of chemically induced lymphomas reported in RI studies are also observed in older untreated rats in the RI colony (Soffritti et al. 2006b). However, they are rarely diagnosed in untreated F344N rats (NTP 1999, 2013) and those exposed to the same chemicals (Table 3). Conversely, the type of lymphoma (mononuclear cell leukemia) commonly observed in F344 rats is rarely observed in treated or untreated Sprague-Dawley rats from the RI colony (EPL 2011b, c).

Along with diagnostic issues, questions have been raised concerning the RI background rate for lymphoma/leukemia. The spontaneous (control) rate of these tumors in RI Sprague-Dawley rats has been reported to be higher than from other sources (Cruzan 2009). Using the meta-regression technique of Sidick and Jonkman (2005), we performed an analysis of past RI studies (See Supplemental Material, Table S1) and identified a significant association between spontaneous lymphoma/leukemia rates and year of study publication for both males and females ($P < 0.001$). The fraction of RI control groups (male or female) with a lymphoma/leukemia rate higher than 10% has increased from 3/43 in 1988-1989 studies to 18/22 in 2002-2006 studies. Possible explanations for this increase include genetic drift associated with inbreeding of the colony and a more active immune system in the non-pathogen-free RI rats. For example, successive inbreeding of Sprague-Dawley rats with chromosome 11 abnormalities has resulted in increased background levels of lymphoblastic lymphoma/leukemias in a Prague colony (Otová et al. 2002). In general, changing conditions (e.g., in husbandry/housing/diet) and differences in pathology examination procedures over time can also contribute to such differences. Caution should be

taken when comparing study results to historical data that are not proximate to the study in question, with the most relevant data coming from the same laboratory, supplier, and within 2 or 3 years of study date (U.S. EPA 2005a).

Detailed characterization of the agent and administered doses

Guidelines developed by regulatory Agencies such as the EPA (U.S. EPA 1998) provide important considerations regarding the source, chemical characterization and storage of a test substance and its incorporation into feed or another administration. RI published reports do not always provide analytical specifications of test substance purity, details of the exposure protocol, or consumption of the test diet or treated water by the animals (see "Good Laboratory Practices"). RI has indicated on their website (Istituto Ramazzini 2013a) that such information is available upon request, but only for RI studies of aspartame, methanol, MTBE and TAME at this time.

Challenging doses and durations of exposure and observation

Consistent with EPA (U.S. EPA 1998) and NTP (Melnick et al. 2008) recommendations, RI uses at least three dose levels – 1) the maximum tolerated level, 2) a dose within an order of magnitude of human exposure levels, and 3) an intermediate level (Soffritti et al. 2002c). RI performs range-finding studies if maximum tolerated levels are not available from the scientific literature. NTP uses data from pre-chronic or subchronic studies (4-13 weeks duration) to estimate the maximally tolerated dose, or the minimally toxic dose, but also suggests using pharmacokinetic information to ensure that no more than one of the selected doses is above a level that saturates the processes of absorption, metabolic activation, or detoxification (Melnick et al. 2008).

RI cancer bioassays use the same duration of exposure as NTP bioassays (NTP 2006), and for those submitted to EPA (U.S. EPA 1998) or to FDA for regulatory review (Hinton 2000) . Typical NTP carcinogenicity studies expose F344N rats and B6C3F1 mice beginning at 6 weeks of age for 2 years. Typical RI carcinogenicity expose Sprague-Dawley rats beginning at 7-10 weeks of age for 2 years, but for some chemicals such as vinyl chloride (Maltoni and Cotti 1988), vinyl acetate monomer (Maltoni et al. 1997a), ethanol (Soffritti et al. 2002a), and aspartame (Soffritti et al. 2006b), exposures were started *in utero.* This *in utero* exposure study design can markedly increase the sensitivity of a cancer bioassay (Melnick et al. 2008; Soffritti et al. 2008).

The most notable difference between RI and other research laboratories is the duration of observation. NTP cancer bioassays are usually terminated at 2 years and the animals sacrificed for analysis. The 2-year termination: (1) maximizes the number of control and treated animals available at the same age for comparisons of pathology and (2) minimizes late-developing background tumors that may limit the ability to detect chemical-induced effects (Melnick et al. 2008). This standard protocol and design has yielded a large database of results in a relatively short period of time (Huff et al. 2008). However, some concerns with the 2-year study design exist. Exposures occurring near the end of the study have little effect on lifetime cancer risk, but adequate data are not available to adjust for this "wasted dose" effect (U.S. EPA 2005b). Although 80% of all human cancers are late-developing, [i.e., occurring after age 60 (Huff et al. 2008)], the 2-year protocol is about 2/3 of the rat lifespan and does not allow sufficient latency for detection of treatment-related late-developing tumors (Bucher 2002; Huff 1999; Maronpot et al. 2004). For these reasons, extension of the rodent study duration used by NTP has been recommended (e.g., Bucher 2002; Huff 1999; Maronpot et al. 2004).

By contrast, the RI observation period is generally the entire "natural" lifespan of the test animal, allowing for the detection of carcinogenic responses after the 2-year treatment period. This aspect has been important for the detection of later-occurring tumors for a number of chemicals [e.g., benzene (Maltoni et al. 1989), xylenes (Maltoni et al. 1997b), mancozeb (Belpoggi et al. 2002a), vinyl acetate monomer (Maltoni et al. 1997a; Minardi et al. 2002), vinyl chloride (Maltoni and Cotti 1988) and acrylonitrile (Maltoni et al. 1988a)]. The advantages of longer observation are reduced for treatments that produce a strong carcinogenic response within 2 years or low survival beyond 2 years. The advantages of a longer observation period would also be offset if test animals experience early mortality from other factors such as laboratory conditions. However, mean 2-year survival of RI Sprague-Dawley rats has been comparable to NTP Sprague-Dawley rats (Caldwell et al. 2008) and above 40% over the past four decades (Belpoggi 2013).

Sufficient numbers of animals per dose group

A major shortcoming of rodent cancer bioassays is its limited statistical power to estimate the true response rate (Melnick et al. 2008). Power is the probability of detecting an effect (i.e., rejecting the null hypothesis) when an effect exists, and it depends on sample size, background effect rate, and magnitude of the true response (Haseman 1984). This limited power may lead to difficulty interpreting non-significant elevations in cancer incidence. Although for some purposes use of few animals may be sufficient (Hinton 2000), the use of at least 100 rodents (50 males and 50 females) per dose level is recommended for most cancer bioassays (Melnick et al. 2008; U.S. EPA 1998, 2005a). The number of animals in any group should not fall below 50 percent at 15 months in mice and 18 months in rats, or below 25 percent at 18 months in mice and 24 months in rats (U.S. EPA 1998).

In RI cancer bioassays, the number of animals is often higher than the 50 animals per sex per dose group typically used by EPA and NTP. Concurrent RI studies have at times shared controls (Belpoggi et al. 1995; Cruzan 2009), with RI publications indicating that such shared controls have been concurrent with, housed in the same facility as, and age-, strain- and colony-matched to treatment groups (Maltoni et al. 1986; Maltoni et al. 1999; Soffritti et al. 2002c; Soffritti et al. 2006a). EPA testing guidelines require (U.S. EPA 1998), and NTP studies generally use (Melnick et al. 2008), concurrent, matched controls. The lack of matched controls would not necessarily preclude a study from contributing to a chemical's cancer weight-of-evidence determination, particularly if relevant (e.g., for the same strain and/or from the same colony) and proximate (e.g., within 3 years of the study in question) historical control data exist (U.S. EPA 2005a).

The potential confounding of treatment-related effects in RI studies by those of litter (i.e., genetic effects) has been raised, as the RI does not always randomize the assignment of animals to treatment groups but often "…assigns all animals from a given litter to the same treatment group, recording which litter each animal came from" (Bucher 2002). However, the RI recently stated that "The assignment of test animals to dose groups will vary in RI studies according to the experimental protocol and aims of the research" and "In the case of experiments in which exposure begins at 6-8 weeks of age (e.g., BT960, methanol), randomization is performed so as to have no more than one female and one male from each litter in each experimental group" (Knowles 2008). For prenatal exposure experiments, "...randomization is performed on the breeders," but the offspring are not randomized across dose groups in order to "...simulate as much as possible the human situation in which all descendents are part of a population" (Knowles 2008). For this latter experimental design, it may be advisable to treat the breeders as

the affected entities or, preferably, to evaluate the dose-response data using nested models that account for intralitter correlations, or the tendency of littermates to respond similarly to one another relative to the other litters in a dose group (U.S. EPA 2012a).

Multiple dose groups for characterization of dose-response relationships

Estimation of the dose-response relationship is a primary aim of carcinogen risk assessment. In general, confidence in dose-response analyses is increased for studies with additional dose groups, particularly when at least two dose levels have response rates above background (U.S. EPA 2012a). EPA testing guidelines recommend (U.S. EPA 1998), and NTP cancer bioassays generally employ, four dose groups (control, low-, mid- and high-dose group). RI cancer bioassays often employ four dose groups as well (Soffritti et al. 2002c), but have employed as many as seven for larger bioassays such as the one performed for aspartame (Soffritti et al. 2005; Soffritti et al. 2006b).

Complete and peer-reviewed histopathological evaluations

Organ system evaluations are well described for EPA and FDA testing requirements (Hinton 2000; U.S. EPA 1998), and for NTP (NTP 2006) and RI (Soffritti et al. 2002c) cancer bioassays. Although diagnostic criteria have been established for most observable lesions, it is not unusual for pathologists to disagree, especially for lesions that are part of a continuum of progressive change (Melnick et al. 2008). As illustrated by the recent PWG of RI studies, a QA pathologist and PWG panel are often used to resolve diagnostic differences between the study and peer review pathologists (Ward et al. 1995). The Society of Toxicologic Pathologists (1997) recommends this type of process "…to ensure that treatment-related findings are properly identified and consistently diagnosed."

The recent PWG review of RI studies (NTP 2011) represents the most in-depth independent review of RI pathologic findings; other, more limited independent reviews of RI histopathological determinations have been performed within the past ten years (Cesta 2008; Hailey 2004; Malarkey et al. 2010). However, not all toxicology laboratories have implemented such a system of review. For instance, the recent Hamner Institute drinking water study of MTBE did not have a PWG review (Bermudez et al. 2012). While the re-evaluation of pathological diagnoses is not an EPA requirement, EPA Office of Pesticide Programs requires the use of a process similar to the NTP PWG when a re-evaluation is conducted (U.S. EPA 1994).

For any peer review of histopathological diagnoses, tissue preservation and condition can be a limiting factor. Studies conducted by or for NTP involve removal of moribund animals to avoid autolytic tissue destruction and to prevent tissue loss through cannibalism (NTP 2006). Though recent RI studies have involved the sacrifice of moribund animals (Soffritti et al. 2010), RI has historically performed pathological examinations on tissues collected solely after natural death, increasing the potential for autolysis and diagnostic difficulties (Hailey 2004; Malarkey et al. 2010). Although RI's use of ethanol rather than the more common formalin for tissue fixation has been questioned (Cesta 2008), ethanol fixation has been used in RI studies for more than 40 years and continues to be used to avoid the toxic effects of formalin, maintain consistency with biopsies taken from human subjects which are also typically use 70% ethanol fixation, and increase comparability of historical controls (Cesta 2008). Ethanol fixation is also advantageous for molecular profiling (Ahram et al. 2003; Chaurand et al. 2008; Gillespie et al. 2002; Kähler et al. 2010; Knowles 2008; O'Leary et al. 2009). As discussed below under "Future Considerations," ethanol fixation also has advantages for microdissection and clonality assays.

As previously discussed, the 2011 PWG review of RI studies found that the histological quality of RI specimens was good and did not affect their review.

Pairwise comparisons and analyses of trends based on survival-adjusted incidence

Existing cancer guidelines recommend trend tests and pairwise comparison tests for determining whether chance, rather than a treatment-related effect, is a plausible explanation for an apparent increase in tumor incidence (U.S. EPA 2005a). In cases in which early mortality is not a critical problem, the Cochran-Armitage test for trend and the Fisher exact test for pairwise comparisons are often used.

For a study with excessive premature death of animals from incidental or chemical-related causes other than tumors at the organ site of interest, animals that died early may not have been alive long enough to contribute a sufficient time at risk to that study (Melnick et al. 2008; U.S. EPA 2005a). Similarly, chemical-related contributing causes of mortality resulting in differential survival across exposure groups may mask treatment-related trends for the effect of interest. In cases of notable early mortality or differential survival, statistical analyses should incorporate survival adjustments (Haseman 1984; U.S. EPA 2005a).

NTP reports typically contain appropriate statistical analyses, including survival-adjusted tumor rates and trend analyses. Recent RI reports also include pairwise comparison and trend tests, and sometimes survival-adjusted analyses (Soffritti et al. 2010), but earlier RI study reports were lacking or inconsistent in this regard and the subsequent application of such tests may be needed for risk assessment purposes. To conduct survival-adjustment analyses, individual animal survival and tumor response data are needed.

For RI datasets for which survival adjustment is warranted, care must be taken in selecting a suitable approach, particularly for incidental (i.e., non-fatal) tumors, for which traditional survival analysis methods may not be valid. The poly-3 adjustment technique used by NTP for survival-adjustment of 2-year bioassay data (Portier and Bailer 1989) is generally useful for incidental data; however, it has never been validated for lifetime studies such as those conducted by RI (Kissling et al. 2008). Early efforts to apply this technique to lifetime studies used the time of death of the last surviving animal as the study length (T), with the result that even animals living 2 years or more could contribute relatively little weight to the incidence denominator representing the animals at risk; this approach may yield erroneous findings (Gebregziabher and Hoel 2009; Kissling et al. 2008). A re-weighting of the individual survival data for lifetime studies such that T is set at 104 weeks, as for 2-year bioassays, and all animals living for 2 years or more contribute a full weight of 1 to the incidence denominator (Gebregziabher and Hoel 2009; Kissling et al. 2008) would better approximate the application of the technique to the 2-year bioassay situation; however, the impacts of applying this weighting scheme to lifetime studies have not been fully investigated, and any differential survival occurring after 2 years would not be accounted for in the survival adjustment.

Good Laboratory Practices

Studies conducted under GLP provide highly detailed protocols, strictly monitor animal health, and extensively document both measurable and observational results (Lilly et al. 1995; OECD 2007). EPA carcinogen risk assessment guidelines do not require GLP certification for laboratory findings to be considered and warn against excluding findings from studies with "limitations of protocol or conduct." However, GLP, or other detailed quality assurance/quality control procedures, are generally used by government laboratories or laboratories explicitly

providing data to governments. The application of GLP decreases the chances for error in standard toxicology testing, ensures transparency and completeness, and minimizes some potential uncertainties that may arise in the interpretation of results.

Independent reviews (Huff 2002) and available RI documentation (Maltoni et al. 1986) suggest that quality control procedures associated with GLP are in place at RI. Following a tour of the RI laboratory and archives, Malarkey et al. (2010) reported "very organized, clean facilities" and that "standard operating procedures (SOPs), GLP documents and necropsy records were within GLP expectations." However, published documentation for RI bioassays is not as detailed as that available from other institutions such as NTP, and has been limited to information provided for individual chemical bioassay reports in journal publications. Some of these individual RI study reports, such as that for trichloroethylene (Maltoni et al. 1988b), contain more detail than others regarding study design and conduct. Reporting variability across RI bioassays and the lack of a single SOP can lead to uncertainty regarding study details. For example, it is not always clear from the individual study reports whether RI used study-specific, concurrent matched controls or common controls across multiple studies (Cruzan 2009). EFSA (2006) has noted deviations in Organization for Economic Co-operation and Development (OECD) guidelines with respect to the RI aspartame study (e.g., a lack of a complete analysis of the test substance, no clear information on the stability of the substance, a lack of clinical observations, a lack of hematological assays, a lack of serology, and limited histopathology reports). While these details may be recorded internally by the RI, they are not readily available for consideration in published reports. In 2009, RI opened a European Experimental Laboratory (EEL) that has GLP certification by the Italian Minister of Health, the Italian GLP compliance monitoring authority,

which will allow them to conduct studies recognized to be in accordance with OECD guidelines (Istituto Ramazzini 2013b).

FUTURE CONSIDERATIONS

Extrapolation of human health risk from laboratory animal studies generally does not address human variability in health status, diet, life style, genetics and other exposures. Rather than attempt to replicate the human scenario, most animal bioassays aim to standardize these factors in test animals so that the contribution of treatment to toxicity and carcinogenic effects can be more readily observed (Bucher 2002; Melnick et al. 2008; NTP 2006; U.S. EPA 1998). RI has put a greater emphasis than most laboratories on study designs "reproducing as much as possible human exposure scenarios" (Soffritti et al. 2006a), and this has contributed to differences between RI and other laboratories in the husbandry and health of experimental animals. In particular, the retention of test animals until death and use of non-pathogen-free conditions have been noted as concerns (Bailey et al. 2012; Bucher 2002; Cruzan 2009; EFSA 2006; Schoeb et al. 2009; Schoeb and Mcconnell 2011a, b).

Health concerns that have the potential to confound study results either through misdiagnoses or premature mortality warrant special consideration. While analogous end-of-life diseases or infections are common in geriatric humans (Caldwell et al. 2008; Schoeb et al. 2009), an evaluation of studies for which test animals demonstrate symptoms or disease late in life must consider the target organ and pathology of the disease or infection and whether it can mask, mimic or reproduce chemical-related effects. Changing lymphoma/leukemia background levels within the RI colony through several decades and the RI's lifespan protocol complicate interpretation of RI findings and comparisons with other laboratory results. Increased efforts by

RI to maintain the healthy and expediently sacrifice the moribund test animals can help, but may not eliminate these inherent difficulties. Continued efforts to transparently report study protocols and results and the continued cooperation and collaboration between RI and other research centers may alleviate some of the concerns discussed in this article. NTP and EPA have collaborated with RI to make detailed reports of several RI bioassays publically available via the RI website. Future efforts such as the NTP/EPA co-sponsored independent PWG review could help to further clarify issues raised about the conduct of RI experiments and accuracy of pathology diagnoses.

Although Belpoggi et al. (1999) described immunoblastic lymphomas after MTBE treatment as progressing from reactive hyperplastic and dysplastic stages to various degrees of malignancy, difficulty can arise in distinguishing between lymphoid neoplastic and reactive changes in the lung when concurrent inflammatory infiltrates are present. The 2011 PWG review of RI studies illustrates such difficulty, but this problem is not unique, especially at the light microscopic level. Extranodal marginal zone B-cell lymphoma diagnoses within a background of a diffuse inflammatory lymphoid infiltrate may be extremely difficult (D'Antonio et al. 2009). Without further examination of clonality or origin (i.e., T-cell vs. B-cell), such cells may be histologically distinguishable from normal cells via light microscopy but difficult to distinguish from inflammatory infiltrates. A number of studies have used T-cell markers to label lymphoblastic lymphomas in SD rats (Fujii et al. 2008; Otová et al. 2002).

Approaches for distinguishing between non-neoplastic and neoplastic lymphoid tissue have been based on the generally accepted conclusion that the vast majority of lymphoid malignancies are clonal in origin (i.e., malignant cells have the same clonally rearranged immunoglobulin (Ig) and/or T-cell receptors) (van Dongen et al. 2003), whereas reactive lymphoid proliferations

contain no predominant single clone (Yakirevich et al. 2007). The demonstration of the monoclonality of immunoglobulin heavy chain (IgH) gene rearrangement is an indispensable method for the diagnosis of B-cell lymphoma as well as histocytochemical analyses (Orba et al. 2003). Polymerase chain reaction (PCR) has been used to identify clonality, but its reliability often depends on the relative abundance of the cell population in question and can be impacted by sampling errors and large numbers of "contaminating" cells (Fend and Raffeld 2000; Orba et al. 2003). In addition, the presence of reactive lymphocytes can produce false-negative PCR results, especially if DNA from whole tissue is used (Cong 2002). Identification of clonal lymphocytic populations may be difficult in cases with scant cellular infiltrates or with a heterogeneous population of cells (Yakirevich et al. 2007). In the case of RI lung lymphoma analyses, both heterogeneous lymphoma subtypes and inflammatory infiltrates have been noted.

Microdissection techniques have been developed to select single cells or groups of cells from a heterogeneous tissue sample for molecular analyses. Laser capture microdissection (LCM) uses low energy infrared laser pulse to selectively adhere visually targeted cells and tissue fragments to a thermoplastic membrane. This technique has been used to distinguish noncancer and cancer tissues and has been an important tool in lymphoma research of human tissues (Liu 2010). The major requirement for effective LCM is correct identification of cell subpopulations in a complex tissue structure. Cells of interest must be identified morphologically with tissue section review and annotation prior to microdissection (Erickson et al. 2009). Therefore, expertise and experience in identifying the cells of interest is critical. A common problem is suboptimal microscopic visualization because of the absence of a mounting medium and a cover slip (Esposito 2007), making precise dissection of cells with a lack of architectural features, such as lymphoid tissues, almost impossible (Fend and Raffeld 2000). As a result, special stains (e.g.,

immunohistochemistry) are used to highlight the cells for isolation and analysis (Fend and Raffeld 2000). Cell number is also critical, with analyses using low cell numbers (i.e., < 2000 cells) subject to false positives from pseudoclonality (Yakirevich et al. 2007). The use of ethanol fixative by RI gives an advantage for immunhistochemistry and LCM studies of clonality, as alcohols fix the tissues by dehydrating them without creating chemical links (Esposito 2007; Orba et al. 2003). Immunohistochemical identification of RI tissues could be used not only as part of the clonality assays to identify tumors but also to compare lymphoma cell types in other rodent assays and in humans. Most rodent and human precursor lymphoma/leukemia is thought to be of T-cell lineage (Fujii et al. 2008; Lin et al. 2005; Otová et al. 2002). It would be of mechanistic interest to find out if the lymphomas identified by RI are also of T-cell origin.

CONCLUSIONS

After investigating the issues identified by EFSA and others, we conclude that RI bioassay results for cancer endpoints other than respiratory tract lymphoma/leukemia, and inner ear and cranium neoplasms, are generally consistent with those of NTP and other laboratories. Concerns regarding a possible link between respiratory infections and the development of lymphomas have been considered (Caldwell et al. 2008). However, a causal association between infections and lymphomas is less likely than the possibility that RI study results have been misinterpreted due to confounding end-of-life respiratory infections. The 2011 PWG review of RI studies showed that, for lymphomas and leukemias, pathologic determinations via light microscopy evaluations are problematic, especially when confounded by infiltrates from an infectious disease. Such diagnoses may vary and depend on pathologists' judgments, process of review, and criteria for a diagnosis (EPL 2011b, c; NTP 2011). As a result, EPA has decided not to rely on lymphoma and

leukemia data from RI studies in Integrated Risk Information System (IRIS) assessments (U.S. EPA 2012b).

In regard to the evaluation of RI cancer bioassays, we have identified several considerations to support interpretation of RI results, summarized in Table 4, which take into account the unique aspects of RI study design within the context of relevant International guidelines. We have also suggested approaches that may assist in the future conduct and review of RI chronic bioassays. While the PCR and microdissection assays discussed have the potential to help resolve some of these concerns for past RI studies, the future resolution of these diagnostic issues will be influenced by the ability of RI to maintain the GLP certification it has obtained from the Italian Ministry of Health and adhere to OECD guidelines regarding the monitoring and control of infectious agents, including regular serological testing, and diligent sacrifices of moribund test animals.

Although the protocols characteristic of RI studies can cause interpretive challenges, aspects of the RI design, including gestational exposure, lifespan observation, and larger numbers of animals and dose groups, may impart advantages that provide chemical risk assessors with valuable insights for the identification chemical-related neoplasia not obtained from other bioassays. We conclude that RI studies may be informative for health risk assessment when reviewed on a case-by-case basis, with consideration given to the unique aspects and issues discussed here that can impact RI bioassay results and interpretations.

References

Apaja M. 1980. Evaluation of toxicity and carcinogenicity of malonaldehyde: An experimental study in Swiss mice. Acta Universitatis Ouluensis, Series D, Medica 55. Finland:Anat Pathol Microbiol. 8:1-61.

Baan R, Grosse Y, Straif K, Secretan B, El Ghissassi F, Bouvard V, et al. 2009. A review of human carcinogens--part F: Chemical agents and related occupations. Lancet Oncol 10(12):1143-1144.

Bailey LA, Prueitt RL, Rhomberg LR. 2012. Hypothesis-Based Weight-of-Evidence evaluation of methanol as a human carcinogen. Regul Toxicol Pharmacol 62(2):278-291.

Belpoggi F, Soffritti M, Maltoni C. 1999. Immunoblastic lymphomas in Sprague-Dawley rats following exposure to the gasoline oxygenated additives Methyl-Tertiary-Butyl Ether (MTBE) and Ethyl-Tertiary-Butyl Ether (ETBE): early observations on their natural history. Eur J Oncol 4:563-572.

Belpoggi F, Soffritti M, Guarino M, Lambertini L, Cevolani D, Maltoni C. 2002a. Results of long-term experimental studies on the carcinogenicity of Ethylene-bis-Dithiocarbamate (Mancozeb) in rats. Ann N Y Acad Sci 982:123-136.

Belpoggi F, Soffritti M, Minardi F, Bua L, Cattin E, Maltoni C. 2002b. Results of long-term carcinogenicity bioassays on tert-amyl-methyl-ether (TAME) and di-isopropyl-ether (DIPE) in rats. Ann N Y Acad Sci 982:70-86.

Belpoggi F. 2013. Graphs of mean survival and body weight in male and female Sprague-Dawley control rats from the colony of the CMCRC/RI (with email to Jeff Gift, dated 5/30/2013).

Benson JM, Gigliotti AP, March TH, Barr EB, Tibbetts BM, Skipper BJ, et al. 2011. Chronic carcinogenicity study of gasoline vapor condensate (GVC) and GVC containing methyl tertiary-butyl ether in F344 rats. J Toxicol Environ Health A 74(10):638-657.

Bermudez E, Willson G, Parkinson H, Dodd D. 2012. Toxicity of methyl tertiary-butyl ether (MTBE) following exposure of Wistar Rats for 13 weeks or one year via drinking water. J Appl Toxicol 32(9):687-706.

Bucher JR. 2002. The National Toxicology Program rodent bioassay: Designs, interpretations, and scientific contributions. Ann N Y Acad Sci 982:198-207.

Burns KM, Melnick RL. 2012. MTBE: recent carcinogenicity studies. Int J Occup Environ Health 18(1):66-69.

BushyRun. 1992a. Final report, methyl tertiary butyl ether - vapor inhalation oncogenicity study in CD-1 mice, with cover letter dated 10/29/92. BRRC report 91N0013A. Washington, DC:Methyl Tertiary Butyl Ether Task Force. http://www.ntis.gov/search/product.aspx?ABBR=OTS0558685 [accessed 1 June 2013].

BushyRun. 1992b. Final report, methyl tertiary butyl ether - vapor inhalation oncogenicity study in Fischer 344 rats, with cover letter dated 11/19/92. BRRC report 91N0013B. Washington, DC:Methyl Tertiary Butyl Ether Task Force. http://www.ntis.gov/search/product.aspx?ABBR=OTS0558686 [accessed 1 June 2013].

Caldwell J, Jinot J, Devoney D, Gift JS. 2008. Evaluation of evidence for infection as a mode of action for induction of rat lymphoma. Environ Mol Mutagen 49(2):155-164.

Cesta MF. 2008. Report of the pathology review of diagnostic criteria for mouse studies at the European Ramazzini Foundation (ERF). Research Triangle Park, NC:National Institute of Environmental Health Sciences.

Cogliano VJ, Grosse Y, Baan RA, Straif K, Secretan MB, El Ghissassi F. 2005. Meeting report: summary of IARC monographs on formaldehyde, 2-butoxyethanol, and 1-tert-butoxy-2-propanol. Environ Health Perspect 113(9):1205-1208.

Cong P. 2002. In situ localization of follicular lymphoma: description and analysis by laser capture microdissection. Blood 99(9):3376-3382.

Cotti G, Maltoni C, Lefemine G. 1988. Long-term carcinogenicity bioassay on vinylidene chloride administered by inhalation to Sprague-Dawley rats: New results. Ann N Y Acad Sci 534:160-168.

Cruzan G. 2009. Assessment of the cancer potential of methanol. Crit Rev Toxicol 39(4):347-363.

D'Antonio A, Caleo A, Licci S, Addesso M, De Palma M, Boscaino A, et al. 2009. A minute focus of extranodal marginal zone B-cell lymphoma arising in Hashimoto thyroiditis diagnosed with PCR after laser capture microdissection: a case report. Thyroid Res 2(1):9.

Duffy PH, Lewis SM, Mayhugh MA, Trotter RW, Hass BS, Latendresse JR, et al. 2008. Neoplastic pathology in male Sprague-Dawley rats fed AIN-93M diet ad libitum or at restricted intakes. Nutr Res 28(1):36-42.

EFSA. 2006. Opinion of the scientific panel on food additives, flavourings, processing aids and materials in contact with food (AFC) on a request from the commission related to a new long-term carcinogenicity study on aspartame. EFSA J 4(5):1-44.

EFSA. 2009. Updated scientific opinion of the panel on food additives and nutrient sources added to food on a request from the European commission related to the 2nd ERF carcinogenicity study on aspartame, taking into consideration study data submitted by the Ramazzini foundation in February 2009. EFSA J 1015:1-18.

EPL. 2011a. Amended pathology quality assessment review and PWG coordinator's report for lifetime carcinogenicity study of ethyl-tertiary-butyl ether (ETBE) in Sprague-Dawley rats conducted at the cancer research center, European Ramazzinni foundation for oncology and environmental sciences, Bologna, Italy: Chemical No./TDMS No.: C00117/00117-64. Durham, NC:Experimental Pathology Laboratories. http://ntp.niehs.nih.gov/ntp/foia/ETBE_amended_qapwg_report_Redacted_508.pdf [accessed 9 September 2013].

EPL. 2011b. Amended pathology quality assessment review and PWG coordinator's report for lifetime carcinogenicity study of methyl alcohol in Sprague-Dawley rats conducted at the cancer research center, European Ramazzinni foundation for oncology and environmental sciences, Bologna, Italy: Chemical No./TDMS No.: C00117E/00117-68. Durham, NC:Experimental Pathology Laboratories. http://ntp.niehs.nih.gov/ntp/foia/pwgMethylAlcohol_508.pdf [accessed 9 September 2013].

EPL. 2011c. Amended pathology quality assessment review and PWG coordinator's report for lifetime carcinogenicity study of methyl-tertiary-butyl ether (MTBE) in Sprague-Dawley rats conducted at the cancer research center, European Ramazzinni foundation for oncology and environmental sciences, Bologna, Italy: Chemical No./TDMS No.: C00117F/00117-69. Durham, NC:Experimental Pathology Laboratories. http://ntp.niehs.nih.gov/ntp/foia/pwgMTBE_508.pdf [accessed 9 September 2013].

EPL. 2011d. Amended pathology quality assessment review and PWG coordinator's report for lifetime carcinogenicity study of vinyl chloride in Sprague-Dawley rats conducted at the cancer research center, European Ramazzinni foundation for oncology and environmental sciences, Bologna, Italy. Durham, NC:Experimental Pathology Laboratories. http://ntp.niehs.nih.gov/ntp/foia/vinylchloride_amended_qapwg_report_Redacted_508BE.pdf [accessed 9 September 2013].

EPL. 2012. Second amended pathology quality assessment review and PWG coordinator's report for lifetime carcinogenicity study of acrylonitrile in Sprague-Dawley rats conducted at the cancer research center, European Ramazzinni foundation for oncology and environmental sciences, Bologna, Italy. Durham, NC:Experimental Pathology Laboratories. http://ntp.niehs.nih.gov/ntp/foia/acrylonitrile_AMENDED_qapwg_report_Redacted_508.pdf [accessed 9 September 2013].

Erickson HS, Albert PS, Gillespie JW, Rodriguez-Canales J, Marston Linehan W, Pinto PA, et al. 2009. Quantitative RT-PCR gene expression analysis of laser microdissected tissue samples. Nature Protocols (Online) 4(6):902-922.

Esposito G, ed. 2007. Complementary techniques: Laser capture microdissection - Increasing specificity of gene expression profiling of cancer specimens. Austin, TX:Landes Bioscience, 54-65.

Fend F, Raffeld M. 2000. Laser capture microdissection in pathology. J Clin Pathol 53(9):666-672.

Frith CH. 1988. Morphologic classification and incidence of hematopoietic neoplasms in the Sprague-Dawley rat. Toxicol Pathol 16:451-457.

Frith CH, Ward JM, Chandra M. 1993. The morphology, immunohistochemistry, and incidence of hematopoietic neoplasms in mice and rats. Toxicol Pathol 21(2):206-218.

Fujii E, Kimura K, Mizoguchi K, Kato A, Takanashi H, Itoh Z, et al. 2008. Assessment of the carcinogenic potential of mitemcinal (GM-611): increased incidence of malignant lymphoma in a rat carcinogenicity study. Toxicol Appl Pharmacol 228(1):1-7.

Gebregziabher M, Hoel D. 2009. Applications of the Poly-K Statistical Test to Life-Time Cancer Bioassay Studies. Hum Ecol Risk Assess 15(5):858-875.

Hailey R. 2001. Letter from Rick Hailey, LEP: Histological evaluation of lesions in rats from carcinogenesis studies conducted by the Ramazzini Foundation (RF) dated Aug 15, 2001. http://hero.epa.gov/index.cfm?action=reference.details&reference_id=1936535 [accessed 9 September 2013].

Hailey JR. 2004. Lifetime study in rats conducted by the Ramazzini Foundation Pathology Working Group Chairperson's report. RTP, NC:NIEHS. http://hero.epa.gov/index.cfm?action=reference.details&reference_id=89842 [accessed 9 September 2013].

Harris NL, Jaffe ES, Stein H, Banks PM, Chan JKC, Cleary ML, et al. 1994. A revised European-American classification of lymphoid neoplasms: a proposal from the International Lymphoma Study Group. Blood 84:1361-1392.

Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, et al. 1999. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting--Airlie House, Virginia, November 1997. J Clin Oncol 17:3835-3849.

Haseman JK. 1984. Statistical issues in the design, analysis and interpretation of animal carcinogenicity studies. Environ Health Perspect 58:385-392.

Hinton DM. 2000. US FDA "Redbook II"; immunotoxicity testing guidelines and research in immunotoxicity evaluations of food chemicals and new food proteins. Toxicol Pathol 28(3):467-478.

Huff J, ed. 1999. Chemicals associated with tumours of the kidney, urinary bladder and thyroid gland in laboratory rodents from 2000 US National Toxicology Program/National Cancer Institute bioassays for carcinogenicity. Lyon, France:International Agency for Research on Cancer, 211-225. http://apps.who.int/bookorders/anglais/detart1.jsp?sesslan=1&codlan=1&codcol=73&codcch=147.

Huff J. 2002. Chemicals studied and evaluated in long-term carcinogenesis bioassays by both the Ramazzini Foundation and the National Toxicology Program: In tribute to Cesare Maltoni and David Rall. Ann N Y Acad Sci 982:208-230.

Huff J, Jacobson MF, Davis DL. 2008. The limits of two-year bioassay exposure regimens for identifying chemical carcinogens. Environ Health Perspect 116(11):1439-1442.

IARC. 2006. Preamble to the IARC monographs. Lyon, France. http://monographs.iarc.fr/ENG/Preamble/ [accessed 1 June 2013].

Ishii H, Koshimizu T, Usami S, Fujimoto T. 1981. Toxicity of aspartame and its diketopiperazine for Wistar rats by dietary administration for 104 weeks. Toxicology 21(2):91-94.

Istituto Ramazzini. 2013a. Istituto Ramazzini - cooperatia sociale onlus. Available: http://www.ramazzini.it/ [accessed 3 June 2013].

Istituto Ramazzini. 2013b. Istituto Ramazzini - cooperatia sociale onlus; Centro di Saggio Europa, European experimental laboratory. Available: http://www.ramazzini.it/centro_di_saggio/ [accessed 3 June 2013].

Kähler D, Alexander C, H S, Abdullah M, Branscheid D, Lindner B, et al. 2010. Proteomics Out of the Archive: Two-dimensional Electrophoresis and Mass Spectrometry Using HOPE-fixed, Paraffin-embedded Tissues. J Histochem Cytochem 58(3):221-228.

Kamata E, Nakadate M, Uchida O, Ogawa Y, Suzuki S, Kaneko T, et al. 1997. Results of a 28-month chronic inhalation toxicity study of formaldehyde in male Fisher-344 rats. J Toxicol Sci 22(3):239-254.

Kerns WD, Pavkov KL, Donofrio DJ, Gralla EJ, Swenberg JA. 1983. Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. Cancer Res 43(9):4382-4392.

King-Herbert AP, Sills RC, Bucher JR. 2010. Commentary: update on animal models for NTP studies. Toxicol Pathol 38(1):180-181.

Kissling G, Portier C, Huff J. 2008. MtBE and cancer in animals: statistical issues with poly-3 survival adjustments for lifetime studies. Regul Toxicol Pharmacol 50(3):428-429.

Knowles K. 2008. Personal communication from Kathryn Knowles to Jeff Gift.

Lilly MZ, Gold KW, Rogers RR. 1995. Integrating QA principles with basic elements of a research program promotes quality science in a non-GLP research laboratory. In Vitro Toxicol 8:191-198.

Lin CW, Liu TY, Chen SU, Wang KT, Medeiros LJ, Hsu SM. 2005. CD94 1A transcripts characterize lymphoblastic lymphoma/leukemia of immature natural killer cell origin with distinct clinical features. Blood 106(10):3567-3574.

Liu A. 2010. Laser capture microdissection in the tissue biorepository. J Biomol Tech 21(3):120-125.

Malarkey D, Herbert R, Nyska A, Sutphin ME, Pernika K. 2010. Report on visit (4/25/2010 - 4/30/2010) and assessment of the pathology procedures performed at the Ramazzini Institute (RI), Bentivoglio, Italy.Malarkey, D; Herbert, R; Nyska, A; Sutphin, ME; Pernika, K.

Maltoni C, Lefemine G, Cotti G. 1986. Experimental research on trichloroethylene carcinogenesis. Princeton, NJ:Princeton Scientific Publishing.

Maltoni C, Ciliberti A, Cotti G, Perino G. 1988a. Long-term carcinogenicity bioassays on acrylonitrile administered by inhalation and by ingestion to Sprague-Dawley rats. Ann N Y Acad Sci 534(ED.):179-202.

Maltoni C, Cotti G. 1988. Carcinogenicity of vinyl chloride in Sprague-Dawley rats after prenatal and postnatal exposure. Ann N Y Acad Sci 534:145-159.

Maltoni C, Lefemine G, Cotti G, Perino G. 1988b. Long-term carcinogenicity bioassays on trichloroethylene administered by inhalation to Sprague-Dawley rats and Swiss and B6C3F1 mice. Ann N Y Acad Sci 534:316-342.

Maltoni C, Ciliberti A, Cotti G, Conti B, Belpoggi F. 1989. Benzene, an experimental multipotential carcinogen: Results of the long-term bioassays performed at the Bologna Institute of Oncology. Environ Health Perspect 82:109-124.

Maltoni C, Ciliberti A, Lefemine G, Soffritti M. 1997a. Results of a long-term experimental study on the carcinogenicity of vinyl acetate monomer in mice. Ann N Y Acad Sci 837(1):209-238.

Maltoni C, Ciliberti A, Pinto C, Soffritti M, Belpoggi F, Menarini L. 1997b. Results of long-term experimental carcinogenicity studies of the effects of gasoline, correlated fuels, and major gasoline aromatics on rats. Ann N Y Acad Sci 837(1):15-52.

Maltoni C, Soffritti M, Belpoggi F. 1999. The scientific and methodological bases of experimental studies for detecting and quantifying carcinogenic risks. Ann N Y Acad Sci 895:10-26.

Maronpot RR, Flake G, Huff J. 2004. Relevance of animal carcinogenesis findings to human cancer predictions and prevention. Toxicol Pathol 32(Suppl 1):40-48.

McConnell EE, Solleveld HA, Swenberg JA, Boorman GA. 1986. Guidelines for combining neoplasms for evaluation of rodent carcinogenesis studies. J Natl Cancer Inst 76(2):283-289.

Melnick RL, Thayer KA, Bucher JR. 2008. Conflicting views on chemical carcinogenesis arising from the design and evaluation of rodent carcinogenicity studies. Environ Health Perspect 116(1):130-135.

Minardi F, Belpoggi F, Soffritti M, Ciliberti A, Lauriola M, Cattin E, et al. 2002. Results of long-term carcinogenicity bioassay on vinyl acetate monomer in Sprague-Dawley rats. Ann N Y Acad Sci 982:106-122.

Molinary VS. 1984. Preclinical studies of aspartame in nonprimate animals. New York:Marcel Dekker, Inc., 289-306.

NEDO. 1987. Toxicological research of methanol as a fuel for power station: summary report on tests with monkeys, rats and mice. Tokyo, Japan.

NTP. 1982. NTP technical report on the carcinogenesis bioassay of vinylidene chloride (CAS No. 75-35-4) in F344/N rats and B6C3F1/N mice (gavage study). NTP TR 228; NTP-80-82; NIH Pub. No. 82-1784. Research Triangle Park, NC. http://ntp.niehs.nih.gov/ntp/htdocs/LT_rpts/tr228.pdf [accessed 1 June 2013].

NTP. 1990. National Toxicology Program technical report on the toxicology and carcinogenesis studies of toluene (CAS no 108-88-3) in F344/N rats and B6C3F1 mice (inhalation studies). NTP TR 371; NIH Pub. No. 90-2826. Research Triangle Park, NC.

NTP. 1992. NTP technical report on the toxicology and carcinogenesis studies of chlorinated water (CAS Nos. 7782-50-5 and 7681-52-9) and chloraminated water (CAS No. 10599-90-3) (deionized and charcoal-filtered) in F344/N rats and B6C3F1 mice (drinking water studies). NTP TR 392; NIH Pub. No. 92-2847. Research Triangle Park, NC. http://ntp.niehs.nih.gov/ntp/htdocs/LT_rpts/tr392.pdf [accessed 1 June 2013].

NTP. 1999. NTP historical control information for the NIH-07 diet. http://ntp.niehs.nih.gov/?objectid=92E6AAA5-F1F6-975E-71C88528A3E7B315 [accessed 1 June 2013].

NTP. 2005. NTP report on the toxicology studies of aspartame (CAS No. 22839-47-0) in genetically modified (FVB Tg.AC hemizygous) and B6.129-Cdkn2atm1Rdp (N2) deficient mice and carcinogenicity studies of aspartame in genetically modified [B6.129-Trp53tm1Brd (N5) haploinsufficient] mice (feed studies). NTP GMM 1; NIH Publication No. 06-4459. Research Triangle Park, NC.

NTP. 2006. Specifications for the conduct of studies to evaluate the toxic and carcinogenic potential of chemical, biological and physical agents in laboratory animals for the national toxicology program (NTP). RTP, NC. http://ntp.niehs.nih.gov/ntp/Test_Info/FinalNTP_ReproSpecsMay2011_508.pdf [accessed 1 June 2013].

NTP. 2011. Summary report of the national toxicology program and environmental protection agencysponsored review of pathology materials from selected ramazzinI institute rodent cancer bioassays. Research Triangle Park, NC. http://ntp.niehs.nih.gov/NTP/About_NTP/Partnerships/International/SummaryPWG_Report_RI_Bioassays.pdf [accessed 1 June 2013].

NTP. 2013. NTP historical controls for NTP-2000 diet. http://ntp.niehs.nih.gov/?objectid=92E705C7-F1F6-975E-72D23026B1645EB9 [accessed 1 June 2013].

O'Leary TJ, Fowler CB, Evers DL, Mason JT. 2009. Protein fixation and antigen retrieval: chemical studies. 84(5):217-221.

OECD. 2007. Establishment and control of archives that operate in compliance with the principles of GLP. Paris, France:OECD Environment Directorate. http://www.ilnas.public.lu/fr/publications/acreditation/bonnes-pratiques-laboratoire/rapports-application/bpl-15-en.pdf [accessed 1 June 2013].

Orba Y, Tanaka S, Nishihara H, Kawamura N, Itoh T, Shimizu M, et al. 2003. Application of laser capture microdissection to cytologic specimens for the detection of immunoglobulin heavy chain gene rearrangement in patients with malignant lymphoma. Cancer 99(4):198-204.

Otová B, Sladká M, Damoiseaux J, Panczak A, Mandys V, Marinov I. 2002. Relevant animal model of human lymphoblastic leukaemia/lymphoma--spontaneous T-cell lymphomas in an inbred Sprague-Dawley rat strain (SD/Cub). Folia Biol (Krakow Pol) 48(6):213-226.

Portier CJ, Bailer AJ. 1989. Testing for increased carcinogenicity using a survival-adjusted quantal response test. Toxicol Sci 12(4):731-737.

Schoeb TR, McConnell EE, Juliana MM, Davis JK, Davidson MK, And Lindsey, J. R. 2009. Mycoplasma pulmonis and lymphoma in bioassays in rats. Vet Pathol 46(5):952-959.

Schoeb TR, McConnell EE. 2011a. Commentary: further comments on Mycoplasma pulmonis and lymphoma in bioassays of rats. Vet Pathol 48(2):420-426.

Schoeb TR, McConnell EE. 2011b. Mycoplasma pulmonis and lymphoma in a methanol bioassay. Vet Pathol 48(4):903-905.

Sidik K, Jonkman JN. 2005. A note on variance estimation in random effects meta-regression. J Biopharm Stat 15(5):823-838.

Society of Toxicologic Pathology. 1997. Documentation of pathology peer review. Position of the Society of Toxicologic Pathologists. Toxicol Pathol 25(6):655.

Soffritti M, Belpoggi F, Lenzi A, Maltoni C. 1997. Results of long-term carcinogenicity studies of chlorine in rats. Ann N Y Acad Sci 837:189-208.

Soffritti M, Belpoggi F, Cevolani D, Guarino M, Padovani M, Maltoni C, eds. 2002a. Results of long-term experimental studies on the carcinogenicity of methyl alcohol and ethyl alcohol in rats. Bologna, Italy:Ann. N. Y. Acad. Sci., 46-69.

Soffritti M, Belpoggi F, Lambertin L, Lauriola M, Padovani M, Maltoni C. 2002b. Results of long-term experimental studies on the carcinogenicity of formaldehyde. Ann N Y Acad Sci 982:87-105.

Soffritti M, Belpoggi F, Minardi F, Maltoni C. 2002c. Ramazzini Foundation cancer program: History and major projects, life-span carcinogenicity bioassay design, chemicals studied, and results. Ann N Y Acad Sci 982:26-45.

Soffritti M, Belpoggi F, Degli ED, eds. 2006a. Cancer prevention: The lesson from the lab. Bologna, Italy:Bononia University Press, 49-64. http://www.ramazzini.it/ricerca/pdfUpload/Cancer%20Medicine%2049-64_2006.pdf [accessed 1 June 2013].

Soffritti M, Belpoggi F, Degli Esposti D, Lambertini L, Tibaldi T, Rigano A. 2006b. First experimental demonstration of the multipotential carcinogenic effects of aspartame administered in the feed to Sprague-Dawley rats. Environ Health Perspect 114:379-385.

Soffritti M, Belpoggi F, Tibaldi E, Esposti DD, Lauriola M. 2007. Life span exposure to low doses of aspartame beginnig during prenatal life increases cancer effects in rats. Environ Health Perspect 115(9):1293-1297.

Soffritti M, Belpoggi F, Esposti DD, Falcioni L, Bua L. 2008. Consequences of exposure to carcinogens beginning during developmental life. Basic & Clinical Pharmacology & Toxicology Online Pharmacology Online 102(2):118-124.

Soffritti M, Belpoggi F, Manservigi M, Tibaldi E, Lauriola M, Falcioni L, et al. 2010. Aspartame administered in feed, beginning prenatally through life span, induces cancers of the liver and lung in male Swiss mice. Am J Ind Med 53(12):1197-1206.

Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al., eds. 2008. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Geneva, Switzerland:WHO Press. http://apps.who.int/bookorders/anglais/detart1.jsp?codlan=1&codcol=70&codcch=4002 [accessed 1 June 2013].

Tobe M, Naito K, Kurokawa Y. 1989. Chronic toxicity study on formaldehyde administered orally to rats. Toxicology 56(1):79-86.

U.S. EPA. 1994. Pesticide Regulation (PR) Notice 94-5: Requests for Re-considerations of Carcinogenicity Peer Review Decisions Based on Changes in Pathology Diagnoses. 731N94002. http://www.epa.gov/PR_Notices/pr94-5.html [accessed 1 June 2013].

U.S. EPA. 1998. Health effects test guidelines: OPPTS 870.4200 Carcinogenicity. EPA/712/C-98/211. Washington, DC. http://www.epa.gov/ocspp/pubs/frs/publications/Test_Guidelines/series870.htm [accessed 1 June 2013].

U.S. EPA. 2005a. Guidelines for carcinogen risk assessment. EPA/630/P-03/001F. Washington, DC. http://www.epa.gov/cancerguidelines/ [accessed 1 June 2013].

U.S. EPA. 2005b. Supplemental guidance for assessing susceptibility from early-life exposure to carcinogens. EPA/630/R-03/003F. Washington, DC. http://www.epa.gov/cancerguidelines/guidelines-carcinogen-supplement.htm [accessed 1 June 2013].

U.S. EPA. 2006. Science Policy Council Peer Review Handbook 3rd Edition. 100B06002. http://nepis.epa.gov/exe/ZyPURL.cgi?Dockey=P1001XCG.txt [accessed 1 June 2013].

U.S. EPA. 2009. HERO, the Database. Available: http://www.epa.gov/hero [accessed 1 June 2013].

U.S. EPA. 2012a. Benchmark dose technical guidance. EPA/100/R-12/001. Washington, DC. http://www.epa.gov/raf/publications/pdfs/benchmark_dose_guidance.pdf [accessed 1 June 2013].

U.S. EPA. 2012b. Update on Ramazzini Institute data in IRIS assessments. Washington, DC. http://epa.gov/iris/ramazzini.htm [accessed 1 June 2013].

van Dongen JJ, Langerak AW, Brüggemann M, Evans PA, Hummel M, Lavender FL, et al. 2003. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia 17(12):2257-2317.

Ward JM, Hardisty JF, Hailey JR, Streett CS. 1995. Peer review in toxicologic pathology. Toxicol Pathol 23(2):226-234.

Yakirevich E, Jackson CL, Meitner PA, Mackenzie D, Tavares R, Robinson-Bostom L, et al. 2007. Analysis of T-cell clonality using laser capture microdissection and high-resolution microcapillary electrophoresis. J Mol Diagn 9(4):490-497.

Zhang L, Steinmaus C, Eastmond DA, Xin XK, Smith MT. 2009. Formaldehyde exposure and leukemia: A new meta-analysis and potential mechanisms. Mutat Res Rev Mutat Res 681(2-3):150-168.

Table 1. Consistent Tumor Diagnoses [a] between RI and QA or PWG pathologists in full or limited pathology analyses. [b]

Organ	Rat Morphology	Mouse Morphology
Adrenal	Cortex: Adenoma, carcinoma, Medulla: benign and malignant pheochromocytomas	
Bone	Femur, rib, or cranium osteosarcoma (partial) [c]	Osteosarcoma
Brain	Malignant astrocytoma and oligodendroglioma, meningioma, reticular cell, granular cell tumor [d]	
Ear	Squamous cell carcinoma (partial) [e]	
Forestomach	Squamous cell papilloma and carcinoma	
Heart	Benign (RI, myxoma) and malignant schwannoma	
Intestine	Large, colon/rectum: adenoma, Small, jejunum: leiomyosarcoma	
Kidney	Renal tubule: adenoma, carcinoma, and lipoma, Transitional epithelium: papilloma, Miscellaneous: ademato us polyp/fibroma, liposarcoma, hemangioma	
Liver	Hepatocellular: adenoma/carcinoma [f], Biliary tumors: (multiple types) [f], Miscellaneous: fibrosarcoma, hemangiosarcoma, hemangioma	Hepatocellular adenoma/carcinoma, Hepatocholangiocarcinoma
Lymph node	Hemangioma	
Lungs	Alveolar/bronchiolar adenoma/carcinoma, squamous cell carcinoma, hemangioma	Alveolar cell adenoma/ carcinoma
Mammary gland	Adenoma/carcinoma, fibroadenoma [g], fibrolipoma, fibrosarcoma (RI, liposarcoma)	
Multiple organs	Histiocytic sarcoma, mononuclear cell leukemia and malignant lymphoma (partial) [h], adenocarcinoma, mesothelioma, schwannoma, osteosarcoma [c], pheochromocytoma, rhabdomyosarcoma	Lymphoma [h]
Nose	Squamous cell carcinoma	
Oral mucosa	Squamous cell carcinoma and papilloma	
Pancreas	Islet cell adenoma	
Parathyroid	Adenoma	
Peritoneum	Fibrosarcoma (RI, liposarcoma), lipoma	
Peripheral nerve	Schwannoma, paraganglioma	
Pituitary gland	Pars distalis adenoma/carcinoma	
Prostate	Adenocarcinoma	
Skeletal muscle	Rhabdomyosarcoma	
Skin	Fibroma (RI, fibrolipoma), fibrosarcoma (RI, liposarcoma), sarcoma, malignant schwannoma, squamous cell carcinoma, Subcutaneous: hemangioma, subcutaneous hemangiosarcoma	Fibrosarcoma
Spleen	Hemangiosarcoma, hemangioma, leiomyosarcoma	
Testes	Interstitial cell: adenoma, Seminal vesicle: adenocarcinoma	
Thymus	Benign thyoma, hemangioma	
Thyroid gland	C cell: adenoma/carcinoma, Follicular cell: adenoma/carcinoma	
Urinary bladder	Leiomyoma, papilloma	
Uterus	Adenoma/carcinoma, fibromyoma, leiomyosarcoma, leiomyoma, stromal polyp/sarcoma, malignant schwannoma	
Vagina	Polyp, sarcoma, leiomyoma	
Zymbal gland	Squamous cell carcinoma	

[a] Listed tumors represent general agreement between the RI and QA or PWG pathologists in full or limited pathology analyses unless noted as "partial," which indicates significant disagreement with some but not all neoplastic diagnoses.

[b] Data sources were RI, QA and PWG diagnostic comparisons of Methanol (EPL 2011b), MTBE (EPL 2011c), ETBE (EPL 2011a), vinyl chloride (EPL 2011d), and acrylonitrile (EPL 2012) rat data, and more limited or preliminary analyses of mouse (Cesta 2008) or rat data (Hailey 2004; Hailey 2001; Malarkey et al. 2010).

[c] Rare bone osteosarcomas were diagnosed by RI and QA pathologists but with less frequency by QA and PWG pathologists and the femur osteosarcomas were diagnosed as osteosarcoma, skin subcutaneous sarcoma and fibrosarcoma by QA pathologists in rat.

[d] RI pathologists and QA pathologist generally agreed on incidence of primary brain neoplasms in the rat but varied in nomenclature and more specific diagnoses. Diagnoses of meningiomas vs. as granular cell tumors or malignant reticuloses, oligodendrogliomas vs. astrocytomas, and malignant oligodendrogliomas vs. microglioma sometimes differed between RI and QA pathologists.

[e] Although there was general agreement between RI and QA diagnoses of ear squamous cell carcinoma for the MTBE rat study, this was not the case for methanol where a number of the lesions were not considered to be neoplastic by the PWG pathologists.

[f] For liver tumors, updated classification used by QA and PWG pathologists and newer RI studies use hepatocellular adenoma/carcinoma descriptors, with those consisting of hepatocellular and cholangiocellular elements now diagnosed as hepatocholangiomas or hepatocholangiocarcinomas.

[g] RI diagnosed fibroma and fibroadenoma as one type without distinction but QA/PWG pathologists classified them per NTP criteria.

[h] There was consistency in lymphoma/leukemia diagnoses reported in the RI mouse study review (Cesta 2008), but only partial consistency in RI rats studies (especially for methanol). Diagnostic consistency was reported in a limited review of lymphoma subtypes (e.g., lymphocytic, histiocytic, monocytic and/or myeloid origin) in rats (Hailey 2004).

Table 2. Summary Incidences of Malignant Lymphoma[a]/Leukemia[b] Diagnosed by RI, QA and PWG Panel Pathologists in Male and Female Rats of RI MTBE and Methanol Studies.[c]

Dose Group	MTBE			Methanol			
	Control	Low	High	Control	Low	Mid	High
Number of male rats	60	60	60	100	100	100	98-100
RI incidence	8	7	6	26	31	35	39
QA incidence	9	3	3	16	13	14	11
PWG incidence	8	1	3	13	11	13	8
Number of female rats	60	59-60	60	100	100	100	99-100
RI incidence	2	7	12	12	21	22	25
QA incidence	2	6	7	9	6	8	6
PWG incidence	0	1	4	8	4	7	6

RI = RI study diagnosis, QA = QA pathologist, PWG = PWG Panel consensus

[a] "Lymphoma" includes lymphoblastic lymphoma, lymphocytic lymphoma, and lymphoimmunoblastic lymphoma for RI diagnoses. For QA pathologist and PWG, all lymphoma sub types were diagnosed as malignant lymphoma.

[b] "Leukemia" includes lymphoblastic leukemia and myeloid leukemia for RI diagnoses and for QA pathologist and PWG, includes myeloid leukemia and mononuclear cell leukemia.

[c] RI and PWG data were obtained from the PWG summary (NTP 2011); QA data were obtained from the EPL reports (EPL 2011b, c).

Table 3. Comparison of Lymphoma/Leukemia Findings for Studies of Ten Chemicals Identified by RI as Positive for Lymphoma/Leukemia [a]

Chemical	Finding	Lab	Strain	Route	Duration	Source
Aspartame	Positive	RI	SD rat	Diet	Lifespan	(Soffritti et al. 2005; Soffritti et al. 2006b)
	Positive	RI	SD rat	Diet	Gestation-Lifespan	(Soffritti et al. 2007)
	Null	RI	Swiss mice	Diet	Gestation-Lifespan	(Soffritti et al. 2010)
	Null	Non-RI	Wistar rat	Diet	2-Year	(Ishii 1981; Ishii et al. 1981)
	Null	Non-RI	p53 Haploin sufficient mice	Diet	9 mth	(NTP 2005)
	Null	Non-RI	SD rat	Oral	2-yr, 2-Gen	(Molinary 1984)
	Null	Non-RI	mice	Diet	110 weeks	(Molinary 1984)
Chlorinated drinking water	Positive	RI	Female SD rat	Drinking water	Lifespan	(Soffritti et al. 1997) [b]
	Positive	Non-RI	Female F344 rats	Drinking water	2-year	(NTP 1992) [c]
	Null	Non-RI	Male F344 rats; B6C3F1 mice	Drinking water	2-year	(NTP 1992)
DIPE	Positive	RI	SD rat	Gavage	Lifespan	(Belpoggi et al. 2002b)
Formaldehyde	Positive	RI	SD rat	Drinking water	Lifespan	(Soffritti et al. 1989; Soffritti et al. 2002b)
	Null	Non-RI	Wistar rat	Drinking water	2-Year	(Til et al. 1989; Tobe et al. 1989)
	Null	Non-RI	F344 rat; B6C3F1 mice	Inhalation	2-Year	(Battelle 1981; Kerns et al. 1983)
	Null	Non-RI	male F344 rat	Inhalation	28 months	(Kamata et al. 1997) [d]
Mancozeb	Positive	RI	SD rat	Diet	Lifespan	(Belpoggi et al. 2002a)
Methanol	Positive	RI	SD rat	Drinking water	Lifespan	(Soffritti et al. 2002a)
	Positive	Non-RI	Eppley Swiss Webster mice	Drinking water	Lifespan	(Apaja 1980)
	Null	Non-RI	F344 rat, B6C3F1 mice	Inhalation	2-Year	(NEDO 1987)
MTBE	Positive	RI	SD rat	Gavage	Lifespan	(Belpoggi et al. 1995; Belpoggi et al. 1997; Belpoggi et al. 1998; Belpoggi et al. 1999)
	Positive [e]	Non-RI	F344 rat	Inhalation	2-Year	(Benson et al. 2011)
	Null	Non-RI	F344 rat	Inhalation	2-Year	(BushyRun 1992b)
	Null	Non-RI	CD-1 mice	Inhalation	18 months	(BushyRun 1992a)
	Null	Non-RI	Wistar rat	Drinking water	2-Year	(Bermudez et al. 2012)
TAME	Positive	RI	SD rat	Gavage	Lifespan	(Belpoggi et al. 2002b)
Toluene	Positive	RI	SD rat	Gavage	Lifespan	(Maltoni et al. 1997b)
	Null	Non-RI	F344 rat, B6C3F1 mice	Inhalation	2-Year	(NTP 1990) [f]
Vinylidene Chloride	Positive	RI	SD rat	Inhalation	Gestation-Lifespan	(Cotti et al. 1988)
	Null	Non-RI	F344 rat	Gavage	2-Year	(NTP 1982)

[a] Bioassays that performed complete histopathology examinations are included.

[b] Authors state that the increase in lymphomas/leukemias "confirm the results" of NTP (1992) but were "not clearly dose related."

[c] NTP considered the marginal increase in leukemia in female rats to be "equivocal evidence of carcinogenic activity."

[d] A small percentage of the original 32 rats/group survived to 28 months due largely to interim sacrifices at 12, 18 and 24 months.

[e] A positive finding for mononuclear cell leukemia in rats co-exposed to MTBE and gasoline, but not gasoline alone was reported by Burns and Melnick (2012) using data from Benson et al. (2011).

[f] Significant ($P < 0.05$) increases occurred in low-dose female mice: lymphoma (2/48, 9/49, 6/50) and lymphoma or leukemia (7/48, 15/49, 7/50). However, NTP did not consider these increases to be related to exposure because similar effects were not found in the high-dose female mice or in male mice or rats.

Table 4. Potential issues and considerations associated with RI studies

Issue	Considerations
Consistency with other laboratories	Consider possible reasons for inconsistencies between RI results and other laboratories, including genetic drift in the RI-bred animal colonies and study differences such as exposure route and duration, observation period, animal husbandry, species or strain and pathological examination procedures.
	Evaluate each study on a case-by-case basis.
Species/strain sensitivity and use of historical data	Recognize that rodent strains differ in their ability to detect certain types of cancers.
	When a high and/or variable background rate is observed, such as the lymphoma/leukemia background rate in RI colony rats, comparing a study response with historical data can be informative.
	Caution should be taken when historical data are not from the same laboratory or supplier and are more than 3 years removed from the study date (U.S. EPA 2005a).
Chemical purity	If chemical purity is not published, consider contacting RI for this information.
	Identify and rule out impurities as potential causative agents or substances that can interfere with the biological availability of the compound of interest.
Dose levels	Determine the basis for dose levels used in the RI study. Was the MTD based on a precursor study or published studies with a similar study design (e.g., species, strain, exposure regimen)?
	Was the MTD high enough to detect key endpoints?
Lifespan observation and prenatal exposure	Recognize that RI lifespan bioassays, particularly when combined with prenatal exposure, can increase sensitivity for the detection of chemical-related effects.
	Be aware that lifespan studies can result in effects that are difficult to distinguish, and thus under- or over-reported, due to high late-life background pathology.
	Consider using a nested dose-response model (U.S. EPA 2012a) to account for possible intralitter correlations/litter effects in RI prenatal exposure studies.
Early mortality and survival adjustments	In cases of notable early mortality or differential survival across dose groups it is important to account for survival time.
	Take care in selecting a suitable survival-adjustment method for full lifetime studies such as RI studies.
	When possible, obtain individual animal data to perform statistical analyses based on survival-adjusted tumor rates and for time-to-tumor dose-response modeling.
Health of the test animal	Evaluate past RI studies with the understanding that RI may not have maintained the health of study animals as in the NTP protocol.
	If a disease is noted or suspected, such as respiratory infections, investigate the possibility of an association between the disease and other responses.
Quality of pathology slides	Be aware that the quality and availability of tissues for pathology slides may suffer in the RI lifespan protocol due to tissue autolysis.
	Low tissue sample numbers relative to the number of animals exposed can be an indication of problems with obtaining or preparing quality pathology slides. If data for individual tissues exist, it may be possible to verify or rule this out as an issue.
Pathological diagnoses and combined tumor counts	RI studies for which test animals show signs of infection should be evaluated with great care, particularly for lesions of the upper respiratory tract.
	While the future considerations discussed have the potential to help resolve some concerns, RI findings that can be confounded by respiratory infection (e.g., lymphomas/leukemias) may not be reliable for risk assessment purposes (U.S. EPA 2012b).
	Regarding the RI practice of reporting combined tumor counts, such as total malignant tumors, International guidelines recommend only combining lesions of the same cell type (EFSA 2006; McConnell et al. 1986; U.S. EPA 2005a).